UNIVERSITY OF COPENHAGEN

Machine learning as a supportive tool to recognize cardiac arrest in emergency calls

Blomberg, Stig Nikolaj; Folke, Fredrik; Ersbøll, Annette Kjær; Christensen, Helle Collatz; Torp-Pedersen, Christian; Sayre, Michael R.; Counts, Catherine R.; Lippert, Freddy K.

Published in:
Resuscitation

DOI:
10.1016/j.resuscitation.2019.01.015

Publication date:
2019

Document version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Blomberg, S. N., Folke, F., Ersbøll, A. K., Christensen, H. C., Torp-Pedersen, C., Sayre, M. R., ... Lippert, F. K. (2019). Machine learning as a supportive tool to recognize cardiac arrest in emergency calls. *Resuscitation*, *138*, 322-329. https://doi.org/10.1016/j.resuscitation.2019.01.015

RESUSCITATION 138 (2019) 322–329

Available online at www.sciencedirect.com

Resuscitation

journal homepage: www.elsevier.com/locate/resuscitation

Clinical paper

Machine learning as a supportive tool to recognize cardiac arrest in emergency calls

Stig Nikolaj Blomberg[a,b,*], **Fredrik Folke**[a,b,c], **Annette Kjær Ersbøll**[d], **Helle Collatz Christensen**[a], **Christian Torp-Pedersen**[e,f], **Michael R. Sayre**[g], **Catherine R. Counts**[g], **Freddy K. Lippert**[a,b]

[a] Emergency Medical Services Copenhagen, Denmark
[b] Department of Clinical Medicine, University of Copenhagen, Denmark
[c] Department of Cardiology, Gentofte University Hospital, Denmark
[d] National Institute of Public Health, University of Southern Denmark, Denmark
[e] Department of Clinical Epidemiology, Aalborg University Hospital, Denmark
[f] Department of Health Science and Technology, Aalborg University, Denmark
[g] Department of Emergency Medicine, University of Washington, United States

Abstract

Background: Emergency medical dispatchers fail to identify approximately 25% of cases of out of hospital cardiac arrest, thus lose the opportunity to provide the caller instructions in cardiopulmonary resuscitation. We examined whether a machine learning framework could recognize out-of-hospital cardiac arrest from audio files of calls to the emergency medical dispatch center.

Methods: For all incidents responded to by Emergency Medical Dispatch Center Copenhagen in 2014, the associated call was retrieved. A machine learning framework was trained to recognize cardiac arrest from the recorded calls. Sensitivity, specificity, and positive predictive value for recognizing out-of-hospital cardiac arrest were calculated. The performance of the machine learning framework was compared to the actual recognition and time-to-recognition of cardiac arrest by medical dispatchers.

Results: We examined 108,607 emergency calls, of which 918 (0.8%) were out-of-hospital cardiac arrest calls eligible for analysis. Compared with medical dispatchers, the machine learning framework had a significantly higher sensitivity (72.5% vs. 84.1%, $p < 0.001$) with lower specificity (98.8% vs. 97.3%, $p < 0.001$). The machine learning framework had a lower positive predictive value than dispatchers (20.9% vs. 33.0%, $p < 0.001$). Time-to-recognition was significantly shorter for the machine learning framework compared to the dispatchers (median 44 seconds vs. 54 s, $p < 0.001$).

Conclusions: A machine learning framework performed better than emergency medical dispatchers for identifying out-of-hospital cardiac arrest in emergency phone calls. Machine learning may play an important role as a decision support tool for emergency medical dispatchers.

Keywords: Artificial intelligence, Machine learning, Cardiopulmonary resuscitation, Emergency medical services, Out-of-hospital cardiac arrest, Detection time, Dispatch-assisted cardiopulmonary resuscitation

* Corresponding author at: Emergency Medical Services Copenhagen & Department of Clinical Medicine, University of Copenhagen, Telegrafvej 5, 2750, Copenhagen, Denmark.

E-mail address: Stig.Nikolaj.Fasmer.Blomberg@regionh.dk (S.N. Blomberg).

https://doi.org/10.1016/j.resuscitation.2019.01.015

Received 16 October 2018; Received in revised form 9 January 2019; Accepted 11 January 2019

Introduction

More than 600,000 people a year sustain an out of hospital cardiac arrest (OHCA) the United States and Europe combined.[1,2] OHCA is a life-threatening condition that needs to be recognized rapidly by dispatchers and recognition of OHCA by either a bystander or a dispatcher in the emergency medical dispatch center is a prerequisite for initiation of cardiopulmonary resuscitation (CPR). Delivery of CPR before the arrival of emergency medical services improves survival, making medical dispatchers' recognition of the condition during emergency calls pivotal.[3–5] However, because patients in cardiac arrest constitute a small minority of the overall call volume, recognition of cardiac arrest is difficult and approximately one quarter of all OHCAs are not recognized during the initial emergency conversation with the medical dispatcher.[3,6,7] This missed cardiac arrest recognition and subsequent provision of life-saving CPR for OHCAs remains a major challenge.

Previous research has identified barriers to recognition of OHCA.[3,6,8] Improving early recognition is a goal for both the American Heart Association and the Global Resuscitation Alliance.[9–12] This challenge may benefit from a novel approach using machine learning.

Machine learning frameworks have been applied to non-emergency conditions, showing that a computer can assist with clinical decision-making or screening under certain circumstances.[13–16] However, machine learning technologies have not been used to support clinical decision-making in an acute medical context.[17] If machine learning could improve OHCA recognition, a condition typically representing approximately 1% of all emergency calls, it holds the potential for other more frequent time critical incidents such as stroke, acute myocardial infarction or sepsis.

In this study, a machine learning framework was used to recognize OHCA from unedited recordings of emergency calls to an emergency medical dispatch center, and the performance of the machine learning framework was subsequently assessed. The study aim was threefold: to test if a unique machine learning framework could improve the OHCA recognition rate compared with trained dispatchers, to examine if the machine learning framework could recognize OHCA faster than the medically trained dispatchers, and to identify possible caller or patient subgroups that were more prone to bias from the medical dispatchers or machine learning framework.

Methods

Machine learning framework

Emergency telephone calls contain a vast amount of information that the medical dispatcher must decipher to draw a conclusion about the urgency of the patient's condition and the type of response required. Issues such as background noise and confusing or conflicting information further complicate this process. Machine learning is most commonly described as an approach in which a model or a framework of models analyses data and adapts by learning from its mistakes.[18] We used a machine learning framework created by the company Corti (Corti.ai, Denmark). The machine learning framework is a network of several machine learning models performing specific tasks, in this case deciphering a conversation in a similar manner to a medical dispatcher. When an emergency call is analysed in real-time by the machine learning framework, the audio file is processed without any prior editing or transcription and transformed to a textual representation of the call, which is then analysed and outputted as a prediction of cardiac arrest.

The classification of OHCA is an end-of-call binary verdict. Time-to-prediction is a continuum of intervals and requires other means of analysis. Analyses were based on processing of the raw audio file by the machine learning framework and were not based on manually transcribed data. The audio files were not prepared or edited before processing. To teach the machine learning framework we used a dataset containing the actual audio files, labelled for the absence or presence of an OHCA at the time of the call. Part of the dataset was used for training, and another part was used for validation. We performed k-fold cross validation to avoid evaluation on a biased split, which means the framework was not evaluated on the same files it had been trained on.[19]

Study population

We included all emergency calls to the Emergency Medical Dispatch Center Copenhagen serving the Capital Region of Denmark received between January 1, 2014, and December 31, 2014. The Capital Region of Denmark covers 1.8 million inhabitants, and the Emergency Medical Dispatch Center Copenhagen responds to 110,000 incidents annually, of which approximately 1200 (1.0%) are OHCAs. The medical dispatchers receiving the calls are nurses (70%) or paramedics (30%) with 6 weeks of focused training in communication and prioritization of emergency calls.[6] The decision-making process is supported by a criteria-based dispatch protocol for assessing the calls, guiding decisions about the emergency level and determining the appropriate responses. This is a validated standardized criteria-based, nationwide Emergency Medical Dispatch System.[20,21]

For all incidents, the associated dispatch audio recordings were retrieved to create a dataset for both teaching and evaluating the machine learning framework. Cases of OHCA were identified through the Danish Cardiac Arrest Registry[4,22] where bystanders initiated CPR or EMS professionals attempted CPR. We excluded cases where the audio file was damaged, or the call was disconnected (unsuitable for analyses), cases with EMS witnessed OHCA (cardiac arrest occurred after the call), cases in which CPR was initiated prior to the emergency call (cardiac arrest already recognized), and cases where the patient showed signs of obvious death.

Calls where dispatchers erroneously suspected cardiac arrest were identified to obtain predictive values for dispatchers. These false-positive cases were identified by a standardised free-text search in the dispatch system, where the logs of all incidents were scanned for expressions related to OHCA (arrest, automated external defibrillator, CPR and lifeless while excluding phrases such as 'not arrest' etc.).

All emergency call recordings were identified and labelled according to whether the calls concerned an OHCA. Calls concerning OHCA were comprehensively examined by the investigators using a predefined and pilot-tested case report form.[3] Time-to-recognition of OHCA was defined as cases where the dispatcher or the caller expressed the presence of an OHCA or the need to initiate CPR or use an automated external defibrillator. Time-to-recognition was determined as the interval from the time the call was answered until the time when the definition of cardiac arrest recognition was achieved. Calls where ambulance personal observed signs of irreversible death, but resuscitative efforts had been initiated prior to ambulance arrival were labelled as OHCA for training the machine learning framework but excluded from the analysis.

Statistical analysis

The performance of the machine learning framework was compared to the medical dispatchers as a baseline. The machine learning framework generates a binary prediction of either 0 or 1 for OHCA classification, corresponding to the probability of that condition being present in each emergency call. Sensitivity and specificity of the machine learning framework were calculated to characterize performance in respect to the reference standard, which was defined as OHCA arrest validated via the Danish Cardiac Arrest Registry excluding EMS-witnessed incidents and calls with CPR initiated prior to the start of the call.

We calculated median time-to-recognition for all calls with recognition by either medical dispatcher or machine learning framework. Differences in time-to-recognition between the machine learning framework and the dispatcher for paired observations (i.e. calls where both the dispatcher and the machine learning framework recognized OHCA) were compared using students t-test and signed rank test. The analysis of time-to-recognition on paired observations is illustrated as a Bland–Altman plot.[23] This method is used to compare a new measurement technique with an established one. In the Bland–Altman plot, the average time-to-recognition of the paired observations is plotted on the X-axis, whereas the difference between the same observations is plotted on the Y-axis.

Results for time-to-recognition are presented for all observations and for paired observations. The results are presented with corresponding interquartile ranges (IQR).

Univariate logistic regression analyses were performed to identify patient-, setting-, and dispatcher-related predictors of OHCA recognition in calls where the machine learning framework recognized out-of-hospital cardiac arrest. Results are reported as odds ratios [OR] with 95% confidence intervals [CI] and p-Values when appropriate. p Values of less than 0.05 were considered significant for all analyses.

Data management and statistical analyses were performed using Statistical Analysis System, SAS software, version 9.4 (SAS Institute).

Approvals

The study was approved by the Danish Health Authority (3-3013-1289/1), the Danish Data Protection Agency (Journal nr.: PVH-2018-001, I-Suite nr.: 6172), and the regional ethics committee (18005504).

Results

In 2014, the Emergency Medical Dispatch Center Copenhagen responded to 110,518 emergency incidents; and the associated call for each was retrieved. Patients with obvious signs of death were excluded (n = 958), as were calls that were either damaged or disconnected within the first 10 s (n = 724). We identified OHCAs within the Danish Cardiac Arrest Registry, (n = 1,147). OHCAs witnessed by ambulance personnel (n = 126), and cases where CPR had been initiated prior to the start of the call (n = 103) were excluded leaving 918 OHCA calls and 107,689 non-OHCA calls eligible for analysis. (Fig. 1).

The characteristics of the OHCA calls are shown in Table 1. Of the 918 calls for patients with an OHCA, 665 (72.4%) were recognized by the medical dispatcher, whereas 772 (84.1%) were recognized by the machine learning framework ($p < 0.001$).

Among calls recognized by the machine learning framework, 117 calls were not recognized by the medical dispatchers. Regarding patient characteristics, there were only minor differences between these incidents and incidents recognized by medical dispatcher. The patients appeared to be slightly older

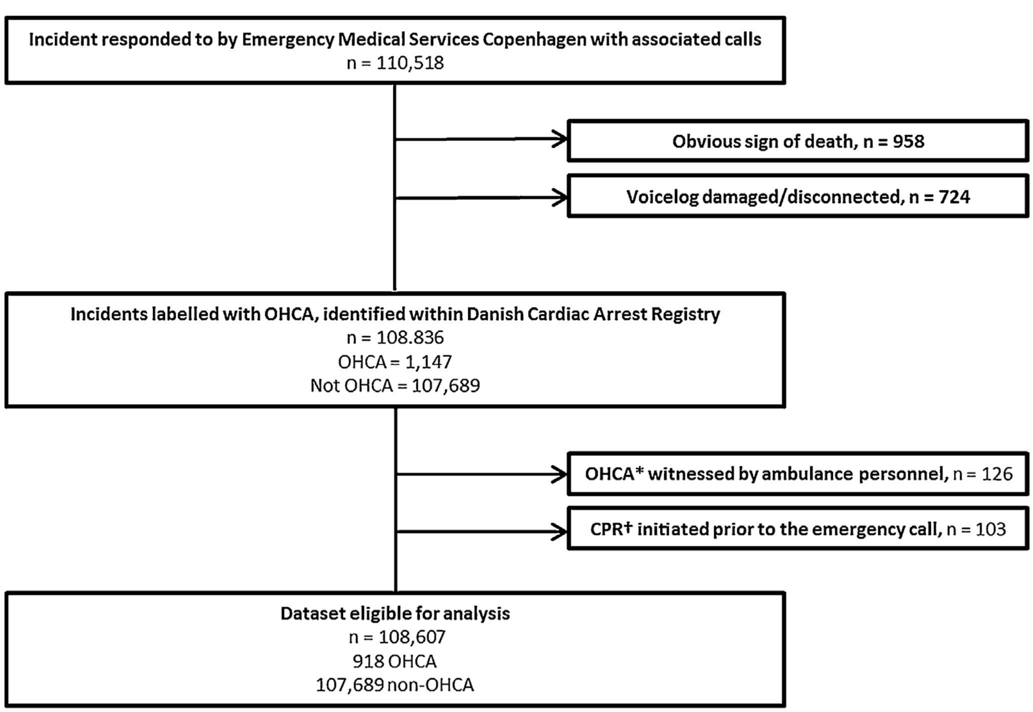

Fig. 1 – Data collection and validation of calls to Emergency Medical Dispatch Center Copenhagen in 2014.
*** Out-of-hospital cardiac arrest.**
† Cardiopulmonary resuscitation.

Table 1 – Characteristics of emergency calls with an out-of-hospital cardiac arrest given by proportion and number of missing values.

	CARDIAC arrest recognized by				
	All (N = 918)	Dispatcher (N = 665)	Machine learning framework (N = 772)	Machine learning framework and not dispatcher (N = 117)	Dispatcher and not machine learning framework (N = 10)
Patient age					
Age, median (Q1–Q3)	71 (61–81)	70 (61–81)	71 (61–81)	74 (67–84)	63 (50–78)
Age missing[a]	29	19	21	4	2
Patient gender					
Male	555 (63.4%)	407 (64.1%)	473 (63.8%)	68 (60.7%)	2 (33.3%)
Female	321 (36.6%)	228 (35.9%)	268 (36.2%)	44 (39.3%)	4 (66.7%)
Missing[a]	42	30	31	5	4
Bystander gender					
Male	349 (38.4%)	248 (37.5%)	281 (36.7%)	40 (34.8%)	7 (70.0%)
Female	561 (61.6%)	414 (62.5%)	486 (63.4%)	75 (65.2%)	3 (30.0%)
N/A[a]	8	3	5	2	0
Bystander (callers) relation to patient					
Caller relative to patient	383 (45.1%)	288 (46.5%)	335 (46.5%)	51 (46.4%)	4 (50.0%)
Caller healthcare professional	238 (28.0%)	169 (27.3%)	198 (27.5%)	31 (28.2%)	2 (25.0%)
Caller all others	228 (26.9%)	162 (26.2%)	189 (26.1%)	28 (25.5%)	2 (25.0%)
N/A[b]	69	46	51	7	2
Access to patient					
Caller by patient's side	654 (78.8%)	508 (83.0%)	571 (80.6%)	71 (67.0%)	8 (88.9%)
Can access patient, but must leave phone	132 (15.9%)	93 (15.2%)	120 (16.8%)	26 (24.5%)	0 (0.0%)
Caller cannot access patient	44 (5.3%)	11 (1.8%)	19 (2.7%)	9 (8.5%)	1 (11.1%)
N/A[b]	88	53	64	11	1
Incident witnessed					
Witnessed by bystander	487 (54.0%)	319 (48.7%)	384 (50.3%)	70 (61.4%)	7 (70.0%)
Not witnessed by bystander	415 (46.0%)	336 (51.3%)	377 (49.7%)	44 (38.6%)	3 (30.0%)
N/A[b]	16	10	13	3	0
Call interrupted					
Call continued until arrival of ambulance	178 (20.5%)	170 (27.8%)	173 (23.5%)	4 (3.6%)	1 (10.0%)
Call ended or interrupted before arrival of ambulance	692 (79.5%)	465 (73.2%)	565 (76.5%)	108 (96.4%)	9 (90.0%)
N/A[b]	48	30	36	5	0
Patient consciousness					
Consciousness addressed	829 (92.3%)	634 (96.2%)	724 (94.8%)	99 (86.8%)	10 (100.0%)
Consciousness not addressed	68 (7.7%)	25 (3.8%)	40 (5.2%)	15 (13.2%)	0
N/A[b]	21	6	10	3	0
Patient breathing					
Breathing addressed	844 (93.2%)	642 (97.0%)	741 (96.4%)	108 (93.1%)	10 (100.0%)
Breathing not addressed	62 (6.8%)	20 (3.0%)	28 (3.7%)	8 (6.9%)	0
N/A[b]	12	3	5	1	0

[a] Missing values are missing civil registration numbers (containing age and sex) with insufficient information on recorded call to identify patient.
[b] N/A is assigned when calls are interrupted or untimely ended.

(74 vs. 70 years) and a slightly less likely to be male (60.7% vs. 64.1%). Comparing this group with calls recognized by medical dispatchers, these calls were less likely to have consciousness addressed by the medical dispatcher (86.8% vs. 96.2%), more likely to be situations where the caller could not access the patient (8.5% vs. 1.8%), and more likely to be bystander witnessed arrests (61.4% vs. 48.7%) than the group of calls recognized by medical dispatchers. Only 10 calls (1.1%) were recognized by the dispatcher but not by the machine learning framework (Table 1).

Among all calls, the machine learning framework reached a sensitivity of 84.1% (95% CI: 81.6–86.3) and a specificity of 97.3% (95% CI: 97.2–97.4) on recognizing OHCA (Table 2). The corresponding sensitivity and specificity of the dispatchers were 72.5% (95% CI: 69.5–75.4) and 98.8% (95% CI:98.7–98.8), respectively. The machine learning framework had a positive predictive value of 21.0% (95% CI: 19.7–22.3) compared with 33.0% (95% CI: 30.1–35.1) for the dispatchers. On all calls recognized by the machine learning model (n = 772), time-to-recognition was significantly shorter for the machine learning framework (median time-to-recognition 44 s, IQR: 24–67)

Table 2 – Characteristics of recognition and time-to-recognition of out-of-hospital cardiac arrests in emergency calls.

Raw audio data for 2014 (n = 108,607)	Machine learning framework	Dispatcher
Sensitivity (95% CI)	84.1 (81.6;86.4)	72.4 (69.4; 75.3)
Specificity (95% CI)	97.3 (97.2;97.4)	98.8 (98.7-98.8)
Negative predictive value (95% CI)	99.9 (99.8;99.9)	99.8 (99.7; 99.8)
Positive predictive value (95% CI)	20.9 (19.6;22.3)	33.0 (30.1; 35.1)
Sensitivity (95% CI), calls unrecognized by dispatchers	44.5 (38.4-50.7)	–
Time-to-recognition, all observations		
Median (95% CI) (seconds)	44 (41; 48)	54 (50; 59)
Lower quartile (seconds)	25	30
Upper quartile (seconds)	72	99
Time-to-recognition, paired observations		
Median (95% CI) (seconds)	41 (38; 44)	54 (50; 59)
Lower quartile (seconds)	24	30
Upper quartile (seconds)	67	97

compared with that of the dispatchers (n = 665) (median time-to-recognition 54 s, IQR: 30–99) ($p < 0.001$). The analysis of paired observations where both the machine learning framework and the dispatcher had recognized the OHCA (n = 655) yielded a median time-to-recognition for the machine learning framework of 44 seconds (IQR: 24–67 s) and for the dispatcher of 54 s (IQR: 30–97 s) ($p < 0.001$).

Time-to-recognition is illustrated in the Bland-Altman plot for paired calls where both the dispatcher and the machine learning framework recognized OHCA (Fig. 2). This plot shows a visual representation that a longer time-to-recognition by the medical dispatcher does not always equate to a longer time-to-recognition by the machine learning framework.

The Bland–Altman plot illustrates that the dispatcher used more time to recognize OHCA than the machine learning framework, as the difference in time-to-recognition (the Y-axis) for the vast majority of calls was greater than zero, which is where there is no difference between the time-to-recognition between dispatcher and machine learning framework. The mean difference was 26 seconds (dashed line, Fig. 2). An observation with a longer mean recognition time on the x axis also has a greater difference in recognition time on the y axis.

Results of univariate logistic regression are shown in Fig. 3 as odds ratios for differences in recognition of OHCA, where OHCAs recognized by the machine learning framework are used as reference. Among calls where the machine learning framework recognized OHCA, dispatchers' recognition was positively associated with the caller's access to the patient. Specifically, when the caller was by the patient's side, the dispatcher was 2.37 times more likely to identify the cardiac arrest than when the caller was not by the patients' side (95% CI 1.50–3.74). If the dispatcher addressed consciousness they had an odds ratio of 3.80 (95% CI 1.90–7.40); and if they addressed breathing, the odds ratio was 2.30 (95% CI 1.00–5.40).

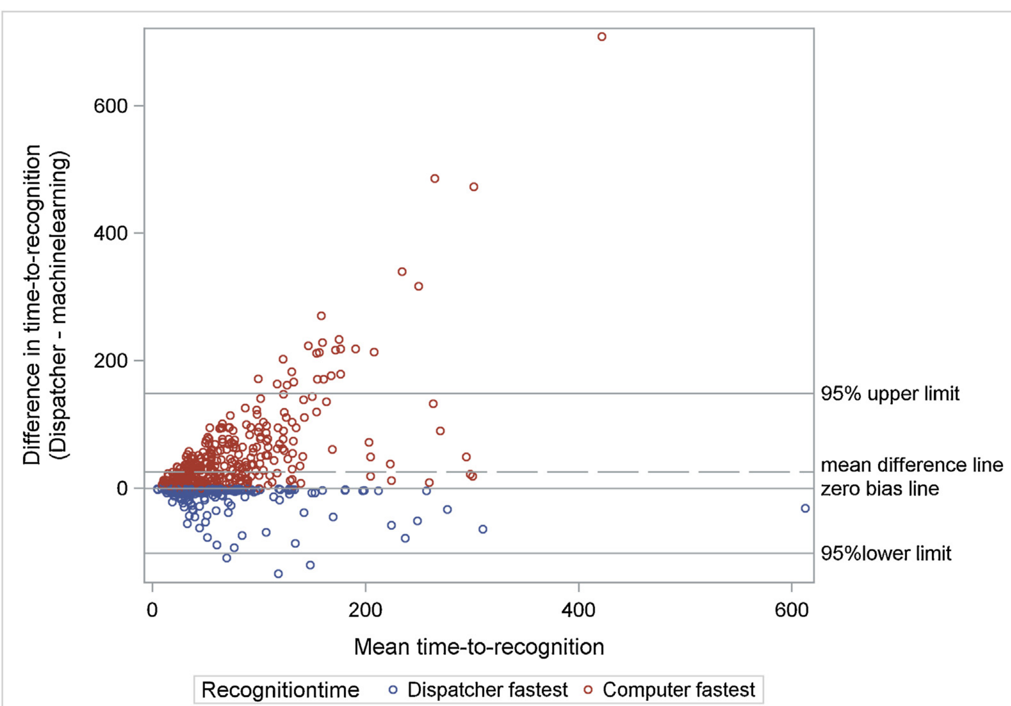

Fig. 2 – Bland-Altman plot comparing time-to-recognition measurements on calls recognized by dispatcher and machine learning framework.

Mean time-to-recognition between medical dispatcher and machine learning framework for each paired observation is plotted on the x-axis while the difference in time-to-recognition for same observation is plotted on the y-axis. Observations where there is no difference in time-to-recognition the observation will be placed along the zero bias line. Observations where the dispatcher recognizes the OHCA faster than the machine learning framework are found below the zero bias line.

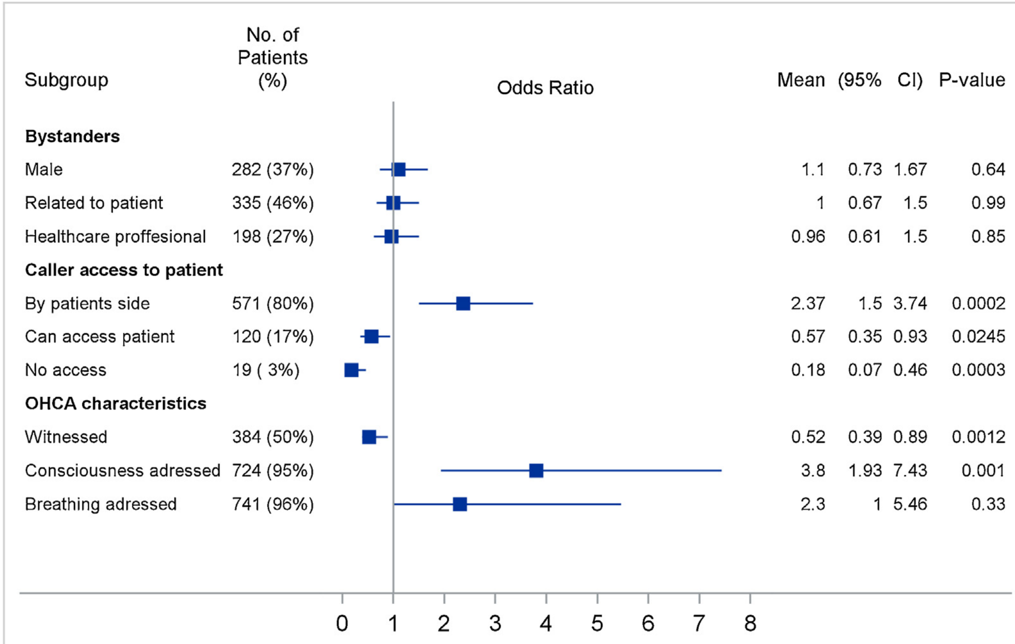

Fig. 3 – Associations between call characteristics and dispatcher recognition of OHCA among calls recognized by machine learning model. (Unadjusted).
The results of regression are shown as odds ratios for differences in recognition of out-of-hospital cardiac arrest (OHCA), where OHCAs recognized by the machine learning framework is used as the reference (vertical line). This means that all observations are recognized by the machine learning framework. An odds ratio greater than one means the factor was positively associated with dispatchers' recognition, and odds ratios less than one mean the factor was associated to dispatcher failing to recognize the OHCA.

When the cardiac arrest was witnessed by a bystander, dispatchers' were less likely to recognise the cardiac arrest compared to the incidents that were unwitnessed (OR 0.52, 95% CI 0.39–0.89) (Fig. 3).

Discussion

The machine learning framework succeeded in recognizing OHCA on raw audio files with a significantly higher sensitivity and similar specificity to the medical dispatchers. It was significantly faster than medical dispatchers in recognizing OHCA. Regression analysis showed the machine learning frameworks overcame some of the barriers to dispatcher recognition that previous studies have identified.[3,6]

Applying a machine learning framework for OHCA holds the potential for increased and faster recognition by dispatchers, increased initiation of telephone CPR efforts, and potentially improving arrest survival. We found a 10 s decrease in recognition time of cardiac arrest. The AHA program guidelines recommend that a high performance system have an elapsed time from call reception to initial dispatch of a response team of less than 60 s, suggesting 120 s be the minimal acceptable standard. In such a perspective 10 s is clinically relevant.[24] These findings open new possibilities for machine learning frameworks, and the potential role they may play as a decision support tool for emergency medical dispatchers in both recognition of OHCA and other time critical conditions.

Machine learning has proven clinically relevant when applied to specific non-urgent medical conditions such as systemic lupus erythematosus or diabetic retinopathy[13–16] showing that machine learning frameworks, though not superior to humans in predicting certain conditions, can support clinicians as a screening tool. However, previous studies analysing images and hospital records were not used in a time-critical setting. In this study the machine learning framework was trained directly on raw, unedited audio files. This is important for the implementation in an acute clinical setting where decision-making has to be efficient and precise. Thus, this approach would make the transition to a live setting easier.

Most studies have published positive predictive values of between 58.4% and 97.9% for professional dispatchers.[25–31] However, the incidence of OHCA in these studies varied greatly, with the studies that reported high positive predictive values also reporting a low incidence of OHCA.[7] Comparatively, the prevalence of OHCA resulted in a positive predictive value of 21.0% by the machine learning framework, meaning that almost four of five machine learning recognized OHCAs would be a false positive. While the positive predictive value of the machine learning framework was lower than the medical dispatcher, a certain amount of over-triage is generally accepted for cardiac arrest and other time-critical incidents.

As such, machine learning should not be used as a stand-alone tool that can independently dispatch ambulances but could act as a supplement to dispatchers' decision-making processes based on standard operating procedures, algorithms and personal experiences. Accordingly, the lower positive predictive value by the machine learning framework should not be a critical failure because it could simply generate cautions about suspected OHCAs and function as an awareness 'flag' for the dispatchers. This could then prompt the dispatchers to increase their focus on presence of breathing and level of consciousness within the ongoing call. In turn, these actions could potentially lead to an increase in the initiation of CPR by bystanders, shown by previous research to improve both short- and long-term survival as well as reducing the risk of anoxic brain damage and nursing-home admission.[4]

Regression analyses illustrated that in the OHCAs recognized by the machine learning framework bystanders' access to the patient along with dispatchers addressing breathing and consciousness were associated with the medical dispatchers' recognition of OHCA on calls also recognized by the machine learning framework. The analysis also illustrated that bystander witnessed OHCAs were negatively associated with recognition of OHCA, which could be explained by the presence of agonal breathing shortly after collapse, which is present in 55% of witnessed OHCAs and may delay or even prevent recognition.[32,33] These findings support those of previous studies.[3]

Limitations

This study has limitations. Predictions by the machine learning framework are made at the termination of the audio recording. In a live setting, the end-of-call prediction is less useful than a predication made while the dispatcher is still on the phone with a bystander. The machine learning framework would need to alert the dispatchers in the case of a suspected OHCA when there is satisfactory confidence in the prediction before the end of the call.

The results from this study need to be tested in another emergency medical setting to prove transmissibility to other languages and organizational cultures. Ideally the use of machine learning should be tested in a randomized controlled trial to measure its impact on patient survival and EMS system operations.

If an OHCA can be recognized from a short conversation over the phone, using machine learning to identify other time critical incidents as stroke, acute myocardial infarct or sepsis holds a great potential. These conditions have a serious health as well as economic impact, and are over twice as prevalent in the United States than OHCAs.[1,34]

Conclusion

Applying a machine learning framework on raw audio-files of emergency calls to identify OHCA showed a significantly higher sensitivity and similar specificity than what was recognized by professional medical dispatchers. Furthermore, the machine learning framework was significantly faster than medical dispatchers in recognizing OHCA albeit with a lower positive predictive value. Machine learning may also play an important role as a decision support tool for emergency medical dispatchers in other time critical conditions.

Conflict of interest

None.

Acknowledgements

We thank Corti: L. Borgholt, M. Singh, A. Wahl-Rasmussen, T.M.S. Tax, L. Maaløe who created the machine learning framework. Corti analysed audio-files labelled according to OHCA for training and testing. Data were subsequently returned with label of prediction for further analysis by the study-team. Corti had no influence on the study's design, results or conclusions. This study was supported by an unrestricted grant from the Danish foundation TrygFonden and The Laerdal Foundation. All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare no support from any organization for the submitted work; no financial relationships with any organizations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

REFERENCES

1. Benjamin EJ, Blaha MJ, Chiuve SE, et al. Heart disease and stroke statistics—2017 update: a report from the American Heart Association. Circulation 2017;135:e146–603.
2. Atwood C, Eisenberg MS, Herlitz J, Rea TD. Incidence of EMS-treated out-of-hospital cardiac arrest in Europe. Resuscitation 2005;67:75–80.
3. Viereck S, Møller TP, Ersbøll AK, et al. Recognising out-of-hospital cardiac arrest during emergency calls increases bystander cardiopulmonary resuscitation and survival. Resuscitation 2017;115:141–7.
4. Kragholm K, Wissenberg M, Mortensen RN, et al. Bystander efforts and 1-year outcomes in out-of-hospital cardiac arrest. N Engl J Med 2017;376:1737–47.
5. Ko SY, Do Shin S, Ro YS, et al. Effect of detection time interval for out-of-hospital cardiac arrest on outcomes in dispatcher-assisted cardiopulmonary resuscitation: a nationwide observational study. Resuscitation 2018;129:61–9.
6. Moller TP, Andrell C, Viereck S, Todorova L, Friberg H, Lippert FK. Recognition of out-of-hospital cardiac arrest by medical dispatchers in emergency medical dispatch centres in two countries. Resuscitation 2016;109:1–8.
7. Viereck S, Moller TP, Rothman JP, Folke F, Lippert FK. Recognition of out-of-hospital cardiac arrest during emergency calls— a systematic review of observational studies. Scand J Trauma Resusc Emerg Med 2017;25:9.
8. Sasson C, Rogers MA, Dahl J, Kellermann AL. Predictors of survival from out-of-hospital cardiac arrest: a systematic review and meta-analysis. Circ Cardiovasc Qual Outcomes 2010;3:63–81.
9. Callaway CW, Donnino MW, Fink EL, et al. Part 8: post-cardiac arrest care: 2015 American Heart Association guidelines update for cardiopulmonary resuscitation and emergency cardiovascular care. Circulation 2015;132:S465–82.
10. Eisenberg M, Lippert F, Castren M, et al. Acting on the call, 2018 update from the Global Resuscitation Alliance. The Global Resuscitation Alliance; 2018. https://www.globalresuscitationalliance.org/wp-content/pdf/acting_on_the_call.pdf.
11. Lippert F, Eisenberg M, Laerdal T, et al. The Global Resuscitation Alliance and Utstein Consensus on implementation of best practice to improve survival from out-of-hospital cardiac arrest. Resuscitation 2018.
12. Nadarajan GD, Tiah L, Ho Ho AFW, et al. Global Resuscitation Alliance Utstein recommendations for developing emergency care systems to improve cardiac arrest survival. Resuscitation 2018;132:85–9.
13. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 2016;316:2402–10.
14. Ceccarelli F, Sciandrone M, Perricone C, et al. Prediction of chronic damage in systemic lupus erythematosus by using machine-learning models. PLoS One 2017;12:e0174200.
15. De Fauw J, Ledsam JR, Romera-Paredes B, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. Nat Med 2018;1.

16. Rajkomar A, Oren E, Chen K, et al. Scalable and accurate deep learning with electronic health records. npj Digital Med 2018;1:18.
17. Stewart J, Sprivulis P, Dwivedi G. Artificial intelligence and machine learning in emergency medicine. Emerg Med Australas 2018;30(6):870–4.
18. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521:436.
19. Devijver PA, Kittler J. Pattern recognition: a statistical approach. Prentice hall; 1982.
20. Andersen MS, Johnsen SP, Sørensen JN, Jepsen SB, Hansen JB, Christensen EF. Implementing a nationwide criteria-based emergency medical dispatch system: a register-based follow-up study. Scand J Trauma Resusc Emerg Med 2013;21:53.
21. Hardeland C, Olasveengen TM, Lawrence R, et al. Comparison of Medical Priority Dispatch (MPD) and Criteria Based Dispatch (CBD) relating to cardiac arrest calls. Resuscitation 2014;85:612–6.
22. Wissenberg M, Lippert FK, Folke F, et al. Association of national initiatives to improve cardiac arrest management with rates of bystander intervention and patient survival after out-of-hospital cardiac arrest. JAMA 2013;310:1377–84.
23. Bland JM, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;327:307–10.
24. American heart association telephone CPR (T-CPR) program recommendations and performance measures. 2018. . Available on: http://cpr.heart.org/AHAECC/CPRAndECC/ResuscitationScience/TelephoneCPR/UCM_493133_Telephone-CPR-T-CPR.jsp.
25. Berdowski J, Beekhuis F, Zwinderman AH, Tijssen JG, Koster RW. Importance of the first link: description and recognition of an out-of-hospital cardiac arrest in an emergency call. Circulation 2009;119:2096–102.
26. Cairns KJ, Hamilton AJ, Marshall AH, Moore MJ, Adgey AJ, Kee F. The obstacles to maximising the impact of public access defibrillation: an assessment of the dispatch mechanism for out-of-hospital cardiac arrest. Heart 2008;94:349–53.
27. Flynn J, Archer F, Morgans A. Sensitivity and specificity of the medical priority dispatch system in detecting cardiac arrest emergency calls in Melbourne. Prehosp Disaster Med 2006;21:72–6.
28. Garza AG, Gratton MC, Chen JJ, Carlson B. The accuracy of predicting cardiac arrest by emergency medical services dispatchers: the calling party effect. Acad Emerg Med 2003;10:955–60.
29. Ma MH-M, Lu T-C, Ng JC-S, et al. Evaluation of emergency medical dispatch in out-of-hospital cardiac arrest in Taipei. Resuscitation 2007;73:236–45.
30. Nurmi J, Pettilä V, Biber B, Kuisma M, Komulainen R, Castrén M. Effect of protocol compliance to cardiac arrest identification by emergency medical dispatchers. Resuscitation 2006;70:463–9.
31. Vaillancourt C, Charette M, Kasaboski A, et al. Cardiac arrest diagnostic accuracy of 9-1-1 dispatchers: a prospective multi-center study. Resuscitation 2015;90:116–20.
32. Bobrow BJ, Zuercher M, Ewy GA, et al. Gasping during cardiac arrest in humans is frequent and associated with improved Survival. Circulation 2008;118(24):2550–4.
33. Clark JJ, Larsen MP, Culley LL, Graves JR, Eisenberg MS. Incidence of agonal respirations in sudden cardiac arrest. Ann Emerg Med 1992;21:1464–7.
34. Lagu T, Rothberg MB, Shieh M-S, Pekow PS, Steingrub JS, Lindenauer PK. Hospitalizations, costs, and outcomes of severe sepsis in the United States 2003–2007. Crit Care Med 2012;40:754–61.